

EXHIBIT A

**UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA
SAN FRANCISCO DIVISION**

RICHARD KADREY, et al.,

Individual and Representative Plaintiffs,

v.

META PLATFORMS, INC., a Delaware
corporation;

Defendant.

Case No. 3:23-cv-03417-VC

**[PROPOSED] STIPULATED ORDER
RE: DISCOVERY OF
ELECTRONICALLY STORED
INFORMATION**

Dept: Courtroom 4 – 17th Floor

Judge: Vince Chhabria

Trial Date: None

Date Action Filed: July 7, 2023

Upon the stipulation of the parties, the Court ORDERS as follows:

1. This Order supplements all other discovery rules and orders. It streamlines Electronically Stored Information (“ESI”) production to promote a “just, speedy, and inexpensive determination of this action, as required by Federal Rule of Civil Procedure 1.”
2. This Order may be modified at the Court’s discretion or by stipulation.

3. As in all cases, costs may be shifted for disproportionate ESI production requests pursuant to Federal Rule of Civil Procedure 26. Likewise, a party's nonresponsive or dilatory discovery tactics are cost-shifting considerations.

4. A party's meaningful compliance with this Order and efforts to promote efficiency and reduce costs will be considered in cost-shifting determinations.

5. The parties are expected to comply with the District's E-Discovery Guidelines ("Guidelines") and are encouraged to employ the District's Checklist for Rule 26(f) Meet and Confer regarding Electronically Stored Information.

6. Pursuant to Federal Rules of Civil Procedure 26(a) and 26(g), and to the extent not already disclosed, the parties shall disclose the following information.

- a. **Custodians:** Each party will identify the ten (10) custodians believed most likely to have discoverable ESI in their possession, custody, or control (hereinafter, "Document Custodians"). The Document Custodians shall be identified by name, title, and connection to the instant litigation. This process shall begin within 14 days of entry of this order, and such identification shall be complete no later than 30 days from entry of this order.
- b. **Additional Custodians:** If, after the Parties identify Document Custodians, a requesting Party believes that additional Document Custodians should be added, then the requesting Party shall advise the producing Party in writing of the proposed additional Document Custodians and the basis for the request. If the Parties have not agreed whether to add the Document Custodian within 30 days of the requesting Party's request, then the matter may be brought to the Court in accordance

with the procedures required by Magistrate Judge Hixson. Nothing in this paragraph waives any objection by a Producing Party that further investigation, discovery or document production is needed before determining the necessity or lack thereof of additional custodians.

- c. **Data Sources:** The Producing Party will inform the Requesting Party about which data sources within the Producing Party's possession, custody, or control have been searched and/or collected from to identify potentially responsive or relevant documents or ESI.
- d. The Parties agree to meet and confer regarding the identification of any additional data sources with potentially responsive information, consistent with the Guidelines for the Discovery of Electronically Stored Information (ESI Guidelines) for the Northern District of California.

7. The Parties agree that each Producing Party is best situated to determine the most appropriate method or methods for that Producing Party to search, collect, cull, and produce documents responsive to discovery, consistent with Sedona Conference Principle 6 which instructs that "[r]esponding parties are best situated to evaluate the procedures, methodologies, and technologies appropriate for preserving and producing their own electronically stored information."

- a. **Search Terms Use:** Each Producing Party may – but is not required to – use search terms to locate potentially relevant documents for collection and/or culling after collection. As a general matter, each Producing Party will use search terms for searching email and other electronic communications, and other sources that are readily amenable to search terms. For other sources, each Producing Party is best equipped to

determine whether search terms or other approaches – such as retrieval of documents and sources identified during custodial interviews – is appropriate for each such source.

- b. **Search Terms:** To the extent search terms are used for a given source or set of data, each Producing Party will develop its own appropriately tailored search terms to locate its relevant documents that are responsive to discovery requests. After applying its search terms to the targeted document set(s), that Producing Party will review a statistically valid (95% \pm 5), randomly generated sample set of the documents that do not hit upon any search terms (“Null Set Sample”) and produce to the Requesting Party the relevant, responsive, non-privileged documents. The Producing Party will develop appropriate new search term(s), if needed, to capture any relevant documents located within the Null Set Sample. The Producing Party will then apply the original and new search terms to the targeted document set and disclose the original and new terms to the Requesting Party, along with a hit report. If the requesting party objects to the sufficiency of the Producing Party’s original and new search terms, the Parties agree to conduct one round (and if there is good cause, i.e. any justifiable reason to do so, more than one round) of the four-step process identified below:

1. The Requesting Party may request one additional Null Set Sample to be reviewed by the Producing Party;
2. The Producing Party will review the additional Null Set Sample and produce all relevant, responsive, non-

privileged documents to the Requesting Party;

3. The Requesting Party can propose additional search terms that would capture the relevant, responsive, non-privileged documents from the additional Null Set Sample; and

4. The Producing Party will apply the search terms to evaluate their reasonableness.

If the search terms proposed by the Requesting Party have an unreasonably high or overbroad yield, the Producing Party may review a randomly generated 95/5 confidence level/margin of error sample set of documents to determine the overbreadth of the proposed search terms.

Where appropriate, the Producing Party may develop alternative search terms that are more narrowly tailored to capture the relevant, responsive, non-privileged documents from the additional Null Set Sample, and provide a hit report on those terms.

As noted above, for good cause, the Requesting Party can request that there be more than one round of the four-step process described above.

For purposes of this provision, “good cause” shall mean any kind of justifiable reason to do the four-step process again.

c. **Other Review Procedures:** Nothing in this Order may be construed or interpreted as precluding a Producing Party from performing a responsiveness review to determine if documents captured by search terms are in fact relevant to the Requesting Party’s discovery requests. Similarly, nothing may be construed or interpreted as precluding a Producing Party

from performing, by any means, a privilege review of documents determined to be relevant. Further, nothing in this Order requires the production of documents that are irrelevant, privileged, or otherwise protected from disclosure merely because the documents hit upon a search term.

- d. **Known, Responsive ESI:** Specific, non-duplicative ESI that is identified by a party as responsive to a discovery request shall not be withheld from review or production solely on the grounds that it was not identified by (or is subject to an exclusion set forth in) the protocols described in, or developed in accordance with, this Order.
- e. **Production and Supplementation of ESI:** Production of responsive, non-privileged documents shall proceed on a rolling basis and be substantially completed in view of the Scheduling Order entered in this case.

8. **Preservation:** To reduce the costs and burdens of preservation and to ensure proper ESI is preserved, the parties agree that:

- A. Parties will preserve non-duplicative, relevant information currently in their possession, custody, or control.
- B. Unwarranted, extraordinary measures will not be taken to preserve the two sources of ESI listed immediately below. Those two sources will be retained pursuant to standard business processes but not searched, reviewed, or produced, unless ordered by the Court or agreed to by the parties:
 - i. backup systems and/or tapes used for disaster recovery; and
 - ii. systems no longer in use that cannot be accessed by using systems currently in use by the party.

C. In addition to the above, the parties agree data from these sources (a) could contain relevant information but (b) under the proportionality factors, should not be preserved (this category will not include sources known to contain data that was used to train any large language model related to issues raised in this litigation or any dataset or records describing that data (hereinafter collectively “LLM Data”), however the parties will meet and confer to determine both any format and reasonable scope regarding the production of any LLM Data relating to Plaintiffs’ copyrighted works):

- i. Deleted, slack, fragmented, or unallocated data only accessible by forensics.
- ii. Random access memory (RAM), temporary files, or other ephemeral data that are difficult to preserve without disabling the operating system.
- iii. On-line data such as temporary internet files, history, cache, cookies, and the like.
- iv. Data in metadata fields that are frequently updated automatically, such as last-opened or last modified dates.
- v. Mobile device activity logs.
- vi. Server, system, or network logs.
- vii. Dynamic fields in databases or log files not stored or retained in the usual course of business.
- viii. Information created or copied during the routine, good-faith performance of processes for the deployment, maintenance, retirement, and/or disposition of computer equipment by the party.

D. The Parties acknowledge that Meta Platforms, Inc. is subject to various privacy

regulations and court orders that require the disposition of identifiable user data.

To the extent that identifiable user data subject to routine disposition as required by privacy regulations and court orders is identified as relevant to the claims or defenses in this case, the parties will meet and confer on preservation of such information as it exists at that time, including the feasibility of such preservation and the type and import of the user data implicated.

- E. Data that is inaccessible within the meaning of Federal Rule of Civil Procedure 26(b)(2)(B) shall not be produced to another party, either in native form or otherwise, unless ordered by the Court. The parties will meet and confer in good faith regarding whether there are reasonable and practical methodologies to query or sample such data sources.

9. **Production Format:** The parties agree to produce documents in the formats described in Exhibit 2 to this Order. If particular documents warrant a different format, the parties will cooperate to arrange for the mutually acceptable production of such documents. The parties agree, to the extent practicable, not to materially degrade the searchability of documents as part of the document production process.

/

/

/

EXHIBIT 1**FIELDS and METADATA TO BE PRODUCED**

1. The load files accompanying scanned paper documents will include the following objective coding fields, to the extent applicable:

Field	Field Description
BEGBATES	First Bates identifier of item
ENDBATES	Last Bates identifier of item
BEGATTACH	Starting Bates number of a document family
ENDATTACH	Ending Bates number of a document family
PAGECOUNT	Number of pages
TEXTLINK	Link to text file for the document
CUSTODIAN	Name of custodian or custodial file
CONFIDENTIALITY	Confidentiality designation assigned to document
REDACTION	Indicate whether document was redacted electronically after scanning

2. The following Metadata fields shall be included in Load files accompanying ESI, to the extent available, and/or unredacted:

Field¹	Field Description for Electronic Documents	Field Description for Emails or Calendar Entries
BEGBATES	First Bates identifier of item	First Bates identifier of item
ENDBATES	Last Bates identifier of item	Last Bates identifier of item

Field¹	Field Description for Electronic Documents	Field Description for Emails or Calendar Entries
BEGATTACH	Starting Bates number of a document family	Starting Bates number of a document family
ENDATTACH	Ending Bates number of a document family	Ending Bates number of a document family
PRODVOL	Production volume	Production volume
PAGECOUNT	Number of pages	Number of pages
ATTACHCOUNT	Number of files attached to parent	Number of files attached to parent
CUSTODIAN	Custodian of the document	Custodian of the document
ALLCUSTODIANS	Individual(s) from whom the document was obtained and de-duplicated out during global de-duplication	Individual(s) from whom the document was obtained and de-duplicated out during global de-duplication
CONFIDENTIALITY	Confidentiality designation	Confidentiality designation

¹ Field Names can vary from system to system and even between different versions of systems. Thus, Parties are to be guided by these Field Names and Field Descriptions when identifying the metadata fields to be produced for a given document pursuant to this ESI Protocol.

Field¹	Field Description for Electronic Documents	Field Description for Emails or Calendar Entries
AUTHOR	Any value populated in the Author field of the document properties	n/a
FROM	n/a	Sender of message or calendar invite
TO	n/a	All recipients that were included on the “To” line of message or calendar invite
CC	n/a	All recipients that were included on the “CC” line of message or calendar invite
BCC	n/a	All recipients that were included on the “BCC” line of message or calendar invite
EMAILSUBJECT	n/a	Subject of message or calendar invite pulled from the document properties
TITLE	Any value populated in the Title field of the document properties	n/a
DATETIMECREATED	Date and time file was created according to filesystem information (format: MM/DD/YYYY HH:mm:ss)	n/a
DATETIMEMODIFIED	Date and time file was last modified according to filesystem information (format: MM/DD/YYYY HH:mm:ss)	n/a
DATETIMESENT	n/a	The sent date and time of the message in the format MM/DD/YYYY HH:mm:ss

Field¹	Field Description for Electronic Documents	Field Description for Emails or Calendar Entries
DATETIMERECEIVED	n/a	The received date of the message in the format MM/DD/YYYY HH:mm:ss
FILENAME	Contents of this Metadata field, or an equivalent	Contents of this Metadata field, or an equivalent
DOCEXT	File extension of document pulled from the document properties	File extension of document pulled from the document properties
FILESIZE	Size of the file in bytes	Size of the file in bytes
FILEPATH	The directory structure of the original file(s). Any container name is included in the path.	The directory structure of the original file(s). Any container name is included in the path.
HASHVALUE	MD5 or SHA1 hash of the document	MD5 or SHA1 hash of the email
THREADID	n/a	Unique identification number that permits threading of email conversations. For instance, unique MS Outlook identification number (“PR_CONVERSATION_INDEX”) is 22 bytes in length, followed by zero or more child blocks each 5 bytes in length, that permits email threading in review software.
ITEMTYPE	Identifies whether the file is an email, attachment, or stand-alone efile.	Identifies whether the file is an email, attachment, or stand-alone efile.

Field¹	Field Description for Electronic Documents	Field Description for Emails or Calendar Entries
NATIVEFILEPATH	Link to native file (if any)	Link to native file (if any)
TEXTFILEPATH	Link to extracted text/OCR file for the document	Link to extracted text/OCR file for the email
REDACTION	Identifies if a document has been redacted	Identifies if a document has been redacted

Such Metadata field values will be provided to the extent available at the time of collection and processing, except that they may be redacted if privileged or if the information in the Metadata is protected by law or Court Order. This list of fields does not create any obligation to create or manually code fields that do not exist as part of the original Metadata of the document.

EXHIBIT 2**PRODUCTION AND LOAD FILE FORMAT FOR DOCUMENTS AND ESI****1. Delimited Text File.**

A delimited text file (.DAT File) containing the fields listed in Exhibit 1 should be provided.

The delimiters for the file should be Concordance defaults accordingly:

Comma	ASCII character 20 (¶)
Quote	ASCII character 254 (þ)
Newline	ASCII character 174 (®)

- The first record should contain the field names in the order of the data.
- All date fields should be produced in mm/dd/yyyy format.
- Use carriage-return line-feed to indicate the start of the next record.
- Load files should not span across media (e.g., CDs, DVDs, Hard Drives, etc.); a separate volume should be created for each piece of media delivered.
- The name of the data load file should mirror the name of the delivery volume, and should have a .DAT extension (i.e., ABC001.DAT).
- The volume names should be consecutive (i.e., ABC001, ABC002, et. seq.).
- If Foreign Language / Unicode text exists, DAT file shall be in appropriate UTF-8 or UTF-16 format.

2. Image Cross-Reference File (.opt): The Opticon cross-reference file is a comma delimited file consisting of six fields per line. There must be a line in the cross-reference file for every image in the database. The format for the file is as follows:

ImageID,VolumeLabel,ImageFilePath,DocumentBreak,FolderBreak,BoxBreak,PageCount.

Further definitions and instructions as follows:

ImageID	The unique designation that Concordance and Opticon use to identify an image. This should be the BegBates Number of the Document.
VolumeLabel	The name of the volume.
ImageFilePath	The full path to the image file.
DocumentBreak	If this field contains the letter “Y,” then this is the first page of a Document. If this field is blank, then this page is not the first page of a Document.
FolderBreak	Leave empty.
BoxBreak	Leave empty.
PageCount	Number of pages in the Document.

- The name of the image load file should mirror the name of the delivery volume, and should have the appropriate extension (e.g., ABC001.LFP).
- The volume names should be consecutive (i.e., ABC001, ABC002, et. seq.).
- There should be one row in the load file per TIFF image.
- Every image in the delivery volume should be contained in the image load file.
- The image key should be named the same as the Bates number of the page.
- Load files should not span across media (e.g., CDs, DVDs, Hard Drives, Etc.), i.e., a separate volume should be created for each piece of media delivered.

3. OCR / Extracted Text Files: OCR or Extracted Text files shall be provided in a separate folder containing Document level text files.

If Foreign Language / Unicode text exists, TEXT files shall be in appropriate UTF-8 or UTF-16 format.

4. TIFFs: Documents that exist only in hard copy format shall be scanned and produced as TIFFs. Documents that exist as ESI shall be converted and produced as TIFFs,

except as provided below. Unless excepted below, single page, black and white, Group IV TIFFs should be provided, at least 300 dots per inch (dpi) for all documents. Each TIFF image shall be named according to a unique corresponding Bates number associated with the document. Each image shall be branded according to the Bates number and the agreed upon confidentiality designation. Original document orientation should be maintained (i.e., portrait to portrait and landscape to landscape). Where the TIFF image is unreadable or has materially degraded the quality of the original, the producing party shall provide a higher quality TIFF image or the native or original file.

5. Color: The parties may request color copies of a limited number of documents where color is necessary to accurately interpret the document.

6. Text Files: A single multi-page text file shall be provided for each document, and the filename should match its respective TIFF filename. When possible, the text of native files should be extracted directly from the native file. Text files will not contain the redacted portions of the documents. A commercially acceptable technology for optical character recognition “OCR” shall be used for all scanned, hard copy documents and for documents with redactions.

7. Native files: Spreadsheet-type files (*e.g.* Microsoft Excel) and PowerPoint presentations will be produced in native format unless redacted, in which instance, spreadsheets will be produced in TIFF with OCR Text Files. To the extent that they are produced in this action, audio, video, and multi-media files will be produced in native format. Native files will be produced with a link in the NATIVEFILEPATH field, along with extracted text (where extracted text is available) and applicable metadata fields set forth in paragraph 4 above. A Bates numbered TIFF placeholder indicating that the document was provided in native format must accompany every native file.

8. Confidentiality Designation: Responsive documents in TIFF format will be stamped with the appropriate confidentiality designations in accordance with the Protective Order entered in this matter. Each responsive document produced in native format will have its confidentiality designation identified in the filename of the native file and indicated on its corresponding TIFF placeholder.

9. Families: Parent-child relationships will be maintained in production. Links within a document are not considered attachments.

IT IS SO STIPULATED, through Counsel of Record.

Dated: April 9, 2024

/s/ Joseph R. Saveri
Counsel for Plaintiffs

Dated: April 9, 2024

/s/ Bobby Ghajar
Counsel for Defendant

IT IS ORDERED that the forgoing Agreement is approved.

Dated:

HON. THOMAS S. HIXSON
UNITED STATES MAGISTRATE JUDGE

SIGNATURE ATTESTATION

Pursuant to Local Rule 5-1(h)(3), I attest under penalty of perjury that concurrence in the filing of this document has been obtained from any other signatory to this document.

Dated: April 9, 2024

By: /s/ Joseph R. Saveri
Joseph R. Saveri